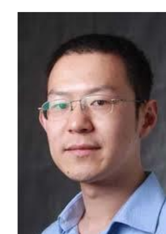
Voting as Maximum Likelihood Estimation: Revisiting the Rationality Assumption

Edith Elkind
University of Oxford









- Elkind, Slinko
  Chapter 8: Rationalizations of Voting Rules,
  in Handbook of Computational Social Choice
  (eds. Brandt, Conitzer, Endriss, Lang, Procaccia)

- Elkind, Shah
  Electing the Most Probable Without
  Eliminating the Irrational:
  Voting Over Intransitive Domains,
  in UAI'14

Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth.

Voting as Preference Aggregation

- What should we get for dinner in Sibiu?
  - Ulle: pork > rabbit > pancake
  - Jerome: rabbit > pancake > pork
  - Edith: pancake > rabbit > pork

- What cake should we order for Yasha's birthday?
  - Dima: chocolate > strawberry > yoghurt
  - Edith: strawberry > yoghurt > chocolate
  - Yasha: yoghurt > strawberry > chocolate

Voting Rules: Plurality

- Plurality: each alternative gets
  - 1 point from each voter who ranks it 1st
  - 0 points from each voter who does not rank it 1st

Ulle: pork > rabbit > pancake
Jerome: rabbit > pancake > pork
Edith: pancake > rabbit > pork

Voting Rules: Borda

- Borda: each alternative gets
  m-i points from each voter who ranks it i^{th}
- Jean-Charles de Borda (1733-1799), French mathematician, physicist, political scientist, sailor

Ulle: pork > rabbit > pancake
Jerome: rabbit > pancake > pork
Edith: pancake > rabbit > pork

Voting Rules: Condorcet

- Condorcet rule:
  - A beats B if majority of voters prefer A to B
  - the Condorcet winner is the alternative that beats every other alternative

- Marquis de Condorcet (1743-1794), French philosopher, mathematician, political scientist

Ulle:	pork > rabbit > pancake
Jerome:	rabbit > pancake > pork
Edith:	pancake > rabbit > pork

Voting Rules: Dodgson

- Dodgson rule:
  - if there is a Condorcet winner, elect him
  - else, try to make each alternative a Condorcet winner by making swaps of adjacent candidates
  - winner: candidate that needs the min # of swaps
- Charles Dodgson, aka Lewis Carroll (1832-1898): English mathematician, logician, Anglican deacon

Voting Rules: Kemeny

- Kemeny rule:
  - for two votes u,v, let $d(u,v) = \# \{(A,B): A >_u B, B >_v A\}$
  - find a ranking that minimizes the total distance to votes
  - output the top alternative is this ranking
- John Kemeny (1926-1992): Hungarian-American mathematician and computer scientist, co-developer of BASIC
  - proposed his rule in 1959

So How Should We Aggregate Preferences?

- Plurality, Borda, Condorcet, Dodgson, Kemeny, …
- Different normative properties
- Choosing a rule may be as hard as choosing the winner!

Ulle: pork > rabbit > pancake
Jerome: rabbit > pancake > pork
Edith: pancake > rabbit > pork

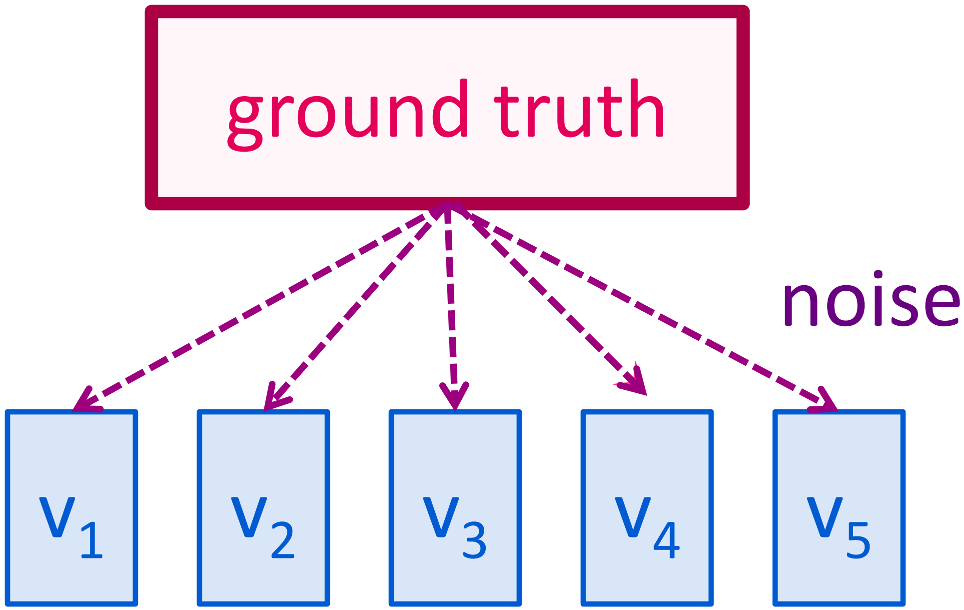
Voting as a Way to Uncover Truth

- Which cleaning company should we hire?
  - Adam: A > B > C
  - Ben: C > B > A
  - Charlie: B > C > A

- Which PhD applicant should we accept?
  - Mike: X > Y > Z
  - Tomasz: Y > X > Z
  - Edith: Z > Y > X

- Medieval church elections
- Crowdsourcing

Voting as Maximum Likelihood Estimation



Which true state of the world is most likely to generate the observed votes?

History

- Marquis de Condorcet (1785), Essai sur l'application de l'analyse a la probabilité des décisions rendues a la pluralité des voix
- H. Peyton Young (1988), Condorcet's theory of voting, *Am. Pol. Sci. Review*
- Conitzer, Sandholm (2005), Common Voting Rules as Maximum Likelihood Estimators, *UAI'05*
- Procaccia, Reddy, Shah (2012), A maximum likelihood approach to selecting sets of alternatives, *UAI'12*
- Elkind, Slinko (forthcoming), Rationalizations of voting rules, *in Handbook of Computational Social Choice*
- Elkind, Shah (2014), Choosing the most probable without eliminating the irrational: voting on intransitive domains, *UAI'14*
- Xia (2014), Note on Young's interpretation of Condorcet's model, *manuscript*

Condorcet -Young's Model

- m alternatives, n voters: $V = (v_1, \ldots, v_n)$
- Ground truth = ranking of the alternatives
- Votes = rankings of the alternatives
- Noise:
  - fix $½ < p < 1$
  - ground truth: u
  - each vote is an outcome of the following process:
    - pick a fresh pair of alternatives a, b; assume $a >_u b$
    - rank them as $a > b$ w.p. p and as $b > a$ w.p. $1-p$
    - if this produces a cycle, restart

Most Likely Ranking [Young'88]

- Kemeny distance: $d(u, v) = |\{(a, b): a >_u b, b >_v a\}|$
- $\phi = p/(1-p)$
- $Pr[v] \sim p^{m(m-1)/2 - d(u, v)} (1-p)^{d(u, v)}$
- $Pr[V] = Pr[v_1] \times \ldots \times Pr[v_n] \sim \phi^{-\Sigma_i d(u, v_i)}$
- $Pr[V] \sim \phi^{-d(u, V)}$
- Most likely ranking: one that minimizes the total distance to votes
  – Kemeny's rule

Rankings vs. Winners

- Finding the most likely ranking: Kemeny's rule
- Finding the most likely winner?
  - Kemeny winners:
    top alternatives of most likely rankings
  - Kemeny winners ≠ most likely winners
- $s_R(a)$: cumulative likelihood of rankings
  where a is ranked first
- $s_R(a) = \Sigma_{u:\, top(u)=a}\, \phi^{-d\,(u,\, V)}$
- Which a maximizes $s_R(a)$?

Most Likely Winner [Y'88, PRS'12]

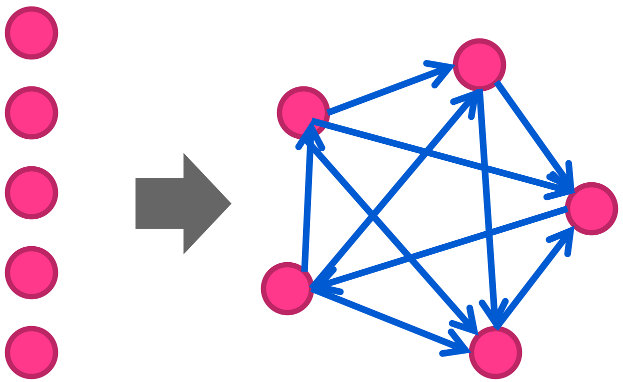
- $s_R(a) = \Sigma_{u:\ top(u)=a}\ \phi^{-d(u,V)}$
- $s_R(a)$: sum of (m-1)! non-positive powers of ϕ
- $p \rightarrow 1$, $\phi = p/(1-p) \rightarrow \infty$ (low noise):
  - the set of most likely winners is a subset of Kemeny winners
- $p \rightarrow 1/2$, $\phi = p/(1-p) \rightarrow 1$ (high noise):
  - the set of most likely winners is a subset of Borda winners

Whodunit?

- Young’88:
  - Condorcet recognized the difference between estimating rankings and estimating winners
  - perhaps he chose not to pursue the winners case because he did not get along with Borda
- Young’88 claims to get Borda for $p \to 1/2$, Maximin (?) for $p \to 1$
  - calculation for an example 3-candidate profile
- PRS’12: more general setting, easily verifiable calculation

Rankings vs. Tournaments

- Working with rankings is hard
  - different pairwise comparisons are not independent
- Can we lower the complexity by allowing non-transitivity…
  - in voters' preferences?
  - in the ground truth?
- I.e., replace rankings with tournaments?



Non-transitive Ground Truth???

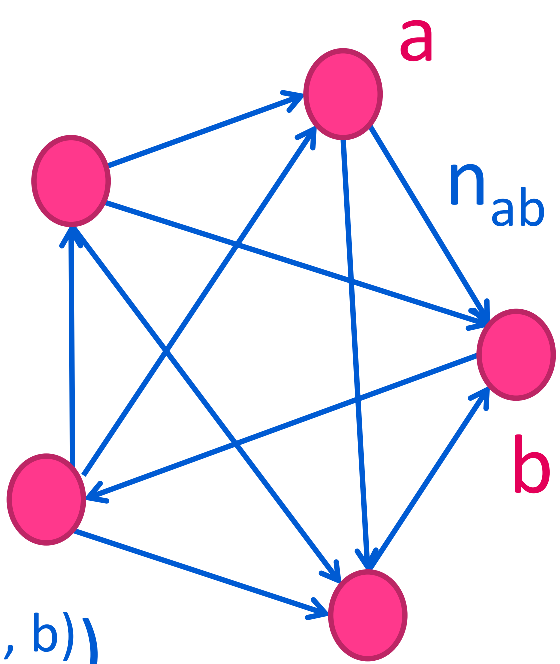
- How can it be that in the true state of the world $A > B$, $B > C$, but $C > A$???
  1. the true state of the world may be obtained by aggregating rankings
  2. just a mental experiment…
- $S_T(a)$ = total likelihood of tournaments where a beats all other alternatives
- most likely winner: $\mathrm{argmax}_a\ S_T(a)$

Most Likely Winner: Tournaments

- $n(a, b)$: # of voters who prefer a to b
- ground truth: T (tournament)
- $\Pr[\, a$ beats everyone in $T \mid V\,]$

$$\sim \prod_{b \neq a} (1+\phi^{\,n(b,a)-n(a,b)})^{-1}$$

- $S_T(a) = \prod_{b \neq a} (1+\phi^{\,n(b,a)-n(a,b)})^{-1}$
- $s_T(a) = 1/S_T(a) = \prod_{b \neq a} (1+\phi^{\,n(b,a)-n(a,b)})$
- most likely alternatives: $\operatorname{argmin}_a s_T(a)$



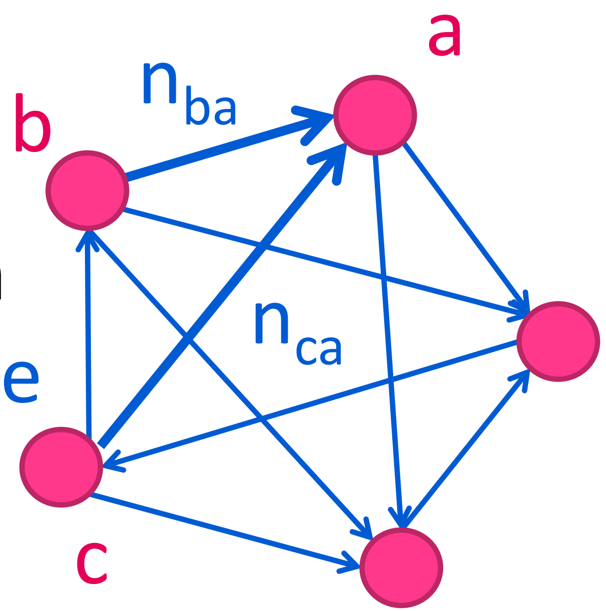
Most Likely Winners: Tournaments

- $s_T(a) = \prod_{b \neq a} (1+\phi^{\,n(b,\,a)\,-\,n(a,\,b)})$
  - most likely alternative: $\min s_T$
- $p \to 1$, $\phi = p/(1-p) \to \infty$ (low noise)
  - most likely winners minimize

    $\sum_{b:\, n(b,\, a) > n(a,\, b)} n(b, a)$ (Tideman's score)
  - hence, the set of most likely winners is a subset of Tideman winners
- $p \to 1/2$, $\phi = p/(1-p) \to 1$ (high noise)
  - the set of most likely winners is a subset of Borda winners

Aside: Tideman's Rule

- $s(a) = \sum_{b:\, n(b,a) > n(a,b)} n(b, a)$
- Proposed in [Tideman'87] as an approximation to Dodgson's rule
- Shown to be an asymptotically optimal approximation to Dodgson's rule [Caragiannis et al., EC'10]
- Elkind, Shah'14: Tideman's rule is a (poly-time) 2-approximation of Kemeny's rule



Whodunit?

- Calculations in [Young'88] are actually for the tournament model!
  - though this is not stated explicitly
- The rule that Young refers to as Maximin is actually Tideman's rule
  - [Tideman'87] vs [Young'88]: communication?
- Young does not point out that for $p \to 1/2$ we get a refinement of Borda, not Borda itself
- [Elkind, Shah'14] clarifying all this (+ more)
- [Xia'14]: same conclusions

Rankings vs. Tournaments

	low noise	**high noise**
rankings	Kemeny*	Borda^{*1}
tournaments	Tideman*	Borda^{*2}

Borda^{*1} and Borda^{*2} may produce disjoint sets of winners [Elkind, Shah'14]

Computational Issues: Rankings

- $s_R(a) = \Sigma_{u:\ top(u)=a}\ \phi^{-d(u, V)}$
  - sum of (m-1)! "polynomial" terms
- $p \rightarrow 1$: computing Kemeny winners is NP-hard
  - even if the winner is unique

  $\Rightarrow$ computing a refinement of Kemeny's rule is NP-hard, too
- $p \rightarrow 1/2$: computing Borda winners is easy
  - but complexity of finding winners in Condorcet-Young's model under high noise is an open problem

Computational Issues: Tournaments

- $s_T(a) = \prod_{b \neq a} (1+\phi^{\,n(b,\,a)\,-\,n(a,\,b)})$
  - "polynomial" of deg $\leq 2mn$
  - product of $m-1$ brackets
  - coefficients can be computed explicitly
- Corollary: for any ϕ (including limit cases), most likely winners are poly-time computable

Agnostic Rule

- For estimating winners, the answer depends on p
  - both for rankings and for tournaments
- Typically, p is not known
- [Elkind, Shah'14]: output all alternatives that are most likely for some p (agnostic rule)
- Experiments: this rule is reasonably decisive

Conclusions and Open Problems

- Even classic papers may sometimes be wrong
  - [Young'88] has 622 citations on Google Scholar
- Writing surveys is useful
- <u>Open</u>: complexity in the rankings model
- <u>Open</u>: understanding the agnostic rule
- <u>Open</u>: other noise models